INTERNATIONAL STANDARD

ISO/IEC 11179-2

First edition
2000-02-15

Information technology — Specification and standardization of data elements —

Part 2: Classification for data elements

Technologies de l'information — Spécification et normalisation des éléments de données —

Partie 2: Classification des éléments de données

Reference number
ISO/IEC 11179-2:2000(E)

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 734 10 79
E-mail copyright@iso.ch
Web www.iso.ch

Printed in Switzerland

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 3.

In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this part of ISO/IEC 11179 may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

International Standard ISO/IEC 11179-2 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information Technology*, Subcommittee SC 32, *Data management and interchange*.

ISO/IEC 11179 consists of the following parts, under the general title *Information technology – Specification and standardization of data elements*:

— *Part 1: Framework for the specification and standardization of data elements*

— *Part 2: Classification for data elements*

— *Part 3: Basic attributes of data elements*

— *Part 4: Rules and guidelines for the formulation of data definitions*

— *Part 5: Naming and identification principles for data elements*

— *Part 6: Registration of data elements*

Introduction

This part of ISO/IEC 11179 provides the basis for documenting, through a specific set of attributes, the classification aspects of data elements. There are many efforts underway to devise classification schemes and to use the schemes to build and populate classification structures. For the purpose of this part of ISO/IEC 11179, the following are all considered types of classification schemes of varying discriminatory power: key words, thesauri, taxonomies, and ontologies. These classification schemes have potentially great utility when associated with various aspects of data elements.

There are several purposes for applying classification to data elements. Classification assists users to find a single data element from among many data elements, facilitates data administration analysis of data elements and, through inheritance, conveys semantic content that is often only incompletely specified by other attributes, such as names and definitions.

The classification schemes accommodated in this Part have utility for

— deriving and formulating abstract and application data elements

— ensuring appropriate attribute and attribute-value inheritance

— deriving names from a controlled vocabulary

— disambiguating

— recognizing superordinate, coordinate, and subordinate data element concepts

— recognizing relationships among data element concepts and data elements

— assisting in the development of modularly designed names and definitions.

Each type of classification scheme mentioned above has particular strengths and weaknesses, and provides the foundation upon which particular capabilities can be built. Keywords, for example, are a quick way to provide users some assistance in locating potentially useful data elements. A thesaurus provides a more structured approach, arranging descriptive terms in a structure of broader, narrower, and related classification categories. A taxonomy provides a classification structure that adds the power of inheritance of meaning from generalized taxa to specialized taxa. Ontologies, with associated epistemologies, can provide rich, rigorously defined structures (e.g., directed acyclic graphs with multiple inheritance) that can convey information needed by software components, such as intelligent agents and mediators, which are useful in the provision of intelligent information services.

The term *data element* refers to *data element type;* the shorter term is used for convenience.

Information technology – Specification and standardization of data elements —

Part 2:
Classification for data elements

1 Scope

This part of ISO/IEC 11179 provides procedures and techniques for associating data with classification schemes. Several components of data elements invite classification—components covered by ISO/IEC 11179—include object classes, properties, representations, value domains, and data element concepts, as well as data elements themselves. The procedures and techniques specified in ISO/IEC 11179-2 will enable Registration Authorities to apply classification schemes that better enable one to

— analyze object classes, data element concepts, and data elements

— make comparisons within the following categories: object classes, properties, representations, data element concepts, and data elements

— reduce the variety of data element concepts and data elements

— identify, describe, and define data element concepts and data elements unambiguously

— assist in the analysis of data elements for the purpose of assigning registration status

— address synonym and homonym problems

— retrieve data element concepts and data elements from a data register

— recognize relationships among data element concepts and data elements

— support the unique and unambiguous identification and referencing of object classes, data element concepts, and data elements in a manner that is linguistically neutral and information technology enabled.

The preparation of ISO/IEC 11179 has also been prompted by the need for standardized data design procedures that will ensure the emergence of data elements capable of supporting electronic data interchange.

This part of ISO/IEC 11179 develops a set of principles, methods, and procedures for specifying what is needed (at a minimum) to document the association between the various components of a data element and one or more classification schemes. This includes the names, nonintelligent identifiers, definitions, and other aspects of the classification scheme and its contents. These can be captured through use of a set of attributes. Particular attributes are specified in this part of ISO/IEC 11179, along with a structure for the contents of these attributes. Users may extend the set of attributes as necessary. Additional information may accompany a taxonomy or ontology; for example, to provide a suggested set of qualifiers that could be applied to the object class, property, or representation taxa to more fully qualify the classification of the particular data element. This part of ISO/IEC 11179 builds upon and utilizes the basic attributes specified in ISO/IEC 11179-3 of ISO/IEC 11179.

An example included in 3.4 shows how selected components of data elements can be associated with a classification scheme through the attributes specified herein. Use of one or more classification schemes is

intended to provide a sound conceptual basis for the development of metadata having enhanced semantic purity and design integrity.

ISO/IEC 11179-2 does not establish a particular classification scheme as preeminent. Sanction of a particular taxonomic approach and/or a particular epistemology is also beyond the scope of ISO/IEC 11179-2. These are addressed by other standards committees and/or tend to be tailored to a particular domain of discourse. The power of the classification scheme and the utility of the content are appropriate areas for competition. Other standards committees are developing or have developed normative languages for use in classification and/or particular techniques and structures that can be accommodated by ISO/IEC 11179-2. For example, the National Information Standards Organization (NISO) has developed a standard for development of a thesaurus. It is appropriate for each classification structure to be documented as to how it was developed and how it can be extended and maintained. Such attributes could be added, by the principle of extensibility, to the attributes specified in ISO/IEC 11179-2. They are not, however, included here.

Each Registration Authority, as described and specified in ISO/IEC 11179-6, may classify components of a data element according to the classification schemes, structures, and content that it deems appropriate. In documenting the classification aspects of data elements, the Registration Authority shall do so according to the principles, methods, procedures, and attributes specified in this part of ISO/IEC 11179.

2 Terms and definitions

For the purposes of this part of ISO/IEC 11179, the following terms and definitions apply.

2.1
attribute
characteristic of an object or entity

2.2
classification scheme
arrangement or division of objects into groups based on characteristics that the objects have in common, e.g., origin, composition, structure, application, and function

2.3
classification scheme item
component of content in a classification scheme

NOTE This may be a node in a taxonomy or ontology, a term in a thesaurus, etc.

2.4
classified component
any component of a data element that may be classified in one or more classification schemes

NOTE The components include the object class, property, representation class, data element concept, value domain, and data element.

2.5
data
representation of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by humans or by automatic means (ISO 2382-4)

2.6
data element
unit of data for which the definition, identification, representation and permissible values are specified by means of a set of attributes

2.7
data element concept
concept that can be represented in the form of a data element, described independently of any particular representation

2.8
name
primary means of identification of objects and concepts for humans

3 Classification attributes for data elements

Attributes shall be utilized to associate various classification schemes with selected components of data elements. These components include object classes, properties, representations, and data element concepts, as well as data elements themselves. As described in ISO/IEC 11179-1, data element concepts are comprised of a relationship between one or more object classes and a property. A data element is formed when a form of representation (e.g., a numeric code) is chosen for the data element concept. Each of the components may be associated with classification schemes as described below.

3.1 Keywords

Keywords are included as basic attributes in ISO/IEC 11179-3. They can be applied to object classes, properties, representations, data elements, and data element concepts. ISO/IEC 11179-3 describes keywords as follows:

```
name        :   Keyword

definition  :   One or more significant words used for retrieval of data
                elements.

obligation  :   Optional

data type   :   Character string

comment     :   This attribute can be used for recording keywords (search keys)
                associated with the data element in question.
```

For the purpose of ISO/IEC 11179-2, “controlled word lists”— wherein each word in a list of words is limited (controlled) to a particular meaning—can be recorded in the same manner as keywords. While keywords are very weak as a classification scheme, they are nonetheless very useful and can be recorded as a “classification scheme type” for use within this standard.

3.2 Thesaurus terms

Thesaurus terms can be associated with data elements and data element concepts. The structure of the thesaurus is not specified in this International Standard. The attributes specified in 3.4 shall be utilized for capturing thesaurus information related to data elements or data element concepts.

3.3 Taxonomy and ontology taxa

A taxonomy is an hierarchical organization of concepts, or taxa, based on generalization/specialization and the mathematical notions of sets, subsets, and set membership. An ontology is a network organization of taxa meant to provide a model of some portion of the world, and consists of theories about the sorts of objects, properties of objects, and relations among objects that are possible in that portion of the world. An ontology may include formal axioms that constrain the interpretation and well-formed use of taxa. The taxa in taxonomies and ontologies may be related to the following classified data registration components: object class, property, representation class, and data element concept. The structure of a taxonomy or an ontology is not specified in ISO/IEC 11179-2, however the attributes listed in 3.4 shall be utilized to capture such structural information.

3.4 Classification attributes

The following attributes may be used for recording the classification information about any component of a data element:

Classified component id

Classified component name

Classification scheme type

Classification scheme name

Classification scheme version

Classification scheme item type

Classification scheme item value

For example:

Classified component id = identifier for object class (e.g., CC0001)

Classified component name = object class (e.g., fruit fly)

Classification scheme type = Taxonomy

Classification scheme name = International Taxonomic Information System (Biological Taxonomy)

Classification scheme version = 2.1

Classification scheme item type = taxon identifier

Classification scheme item value = 5411

or

Classification scheme item type = taxon term

Classification scheme item value = Drosophila

The relationships among the classification attributes are depicted in Figure 1. Additional attributes may be utilized where necessary.

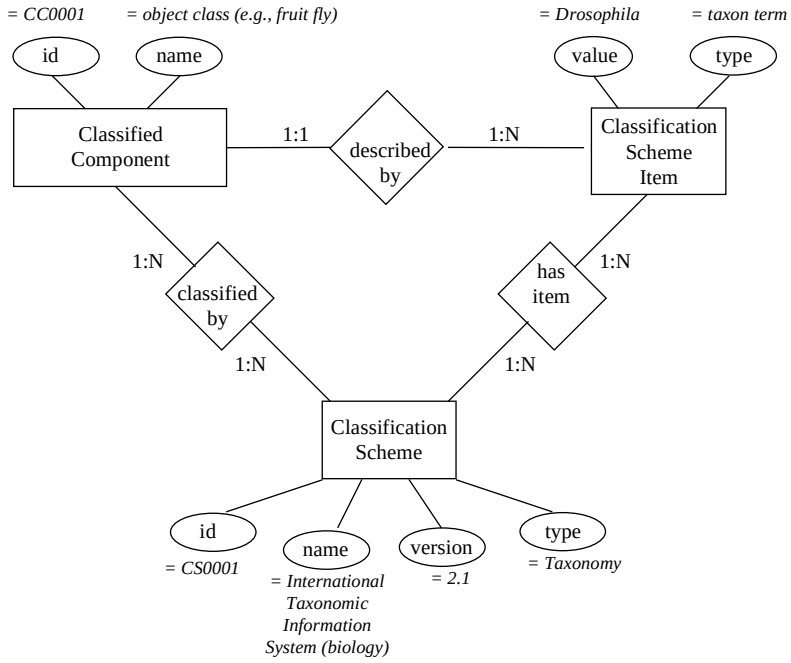

Figure 1 — Classification attributes for administered components, and an example of their usage

3.5 Classification attribute descriptions

3.5.1 name : **Classified component id**

definition : A linguistically neutral, unique, and unambiguous identifier for the identification and referencing of a classified component

obligation : Conditional

condition : Required if the classification items are to be recorded

data type : Character string

comment : Example: CC0001

3.5.2 name : **Classified component name**

definition : The name of the data element component that is subject to classification, such as object class, property, representation class, value domain, data element concept, and data element

obligation : Conditional

condition : Required if the classification items are to be recorded

data type : Character string

comment : Example: object class

3.5.3 name : **Classification scheme type**

definition : The type of classification scheme from which classification items are drawn

obligation : Conditional

condition : Required if the classification items are to be recorded

data type : Character string

comment : Example: Taxonomy

3.5.4 name : **Classification scheme id**

definition : A linguistically neutral, unique, and unambiguous identifier for the identification and referencing of a classified component

obligation : Conditional

condition : Required if the classification items are to be recorded

data type : Character string

comment : Example: CS0001

3.5.5 name : **Classification scheme name**

definition : The name of the particular classification scheme from which the classification items are drawn

obligation : Conditional

condition : Required if the classification items are to be recorded

data type : Character string

comment : Example: Kenworthey's Taxonomy

3.5.6 name : **Classification scheme version**

definition : The version of the particular classification scheme from which the classification items are drawn

obligation : Conditional

condition : Required if the classification items are to be recorded

data type : Character string

comment : Example: “Version 2.1”

3.5.7 name : **Classification scheme item type**

definition : The type of the “classification scheme item value” in 3.5.8.

obligation : Conditional

condition : Required if the classification items are to be recorded.

data type : Character string

comment : Examples: “Unintelligent identifier,” “Term,” or “Definition”

3.5.8 name : **Classification scheme item value**

definition : An instance of the name, or identifier, or definition, etc. of a classification scheme item

obligation : Optional

data type : Character string

comment : Examples: “5411” is an example of a value of type “unintelligent identifier” for a taxon within a taxonomy. “Drosophila” is an example of a value of type “term” used to identify a node in a taxonomy.

Bibliography

[1] ISO/IEC 2382-4, *Information technology — Vocabulary — Part 4: Organization of data*

[2] ISO/IEC 11179-1, *Information technology — Specification and standardization of data elements — Part 1: Framework for the specification and standardization of data elements*

[3] ISO/IEC 11179-3, *Information technology — Specification and standardization of data elements — Part 3: Basic attributes of data elements*

[4] ISO/IEC 11179-6, *Information technology — Specification and standardization of data elements — Part 6: Registration of data elements*

ICS 35.040

Price based on 8 pages